Toward an Architecture for the Global Wordnet Initiative

Andrea Marchetti[1], Maurizio Tesconi[1], Francesco Ronzano[1], Marco Rosella[1] Francesca Bertagna[2], Monica Monachini[2], Claudia Soria[2], Nicoletta Calzolari[2], Chu-Ren Huang[3], Shu-Kai Hsieh[3]

[1]Istituto di Informatica e Telematica-CNR, Via Moruzzi 1, Pisa, Italy
{andrea.marchetti, maurizio.tesconi, francesco.ronzano, marco.rosella}@iit.cnr.it

[2]Istituto di Linguistica Computazionale-CNR, Via Moruzzi 1, Pisa, Italy
{francesca.bertagna, monica.monachini, claudia.soria, nicoletta.calzolari}@ilc.cnr.it,

[3]Academia Sinica, Nankang, Taipei, Taiwan,
churen@gate.sinica.edu.tw, shukai@gmail.com

***Abstract*— Enhancing the development of multilingual lexicons is of foremost importance for intercultural collaboration to take place, as multilingual lexicons are the cornerstone of several multilingual applications. However, the development and maintenance of large-scale, robust multilingual dictionaries is a tantalizing task. Moreover, Semantic Web's growing interest towards the availability of high-quality lexical resources and their multilingual interoperability, is focusing more and more attention on this topic. In this paper we present a tool, based on a web service architecture, enabling semi-automatic generation of bilingual lexicons through linking of distributed monolingual lexical resources. In addition to lexicon development, the architecture also allows enrichment of monolingual source lexicons through exploitation of the semantic information encoded in corresponding entries. In the paper we describe our case study applied to the Italian and Chinese wordnets, and we illustrate how the architecture can be extended to access distributed multilingual WordNets over the Internet, paving the way to exploitation in a cross-lingual framework of the wealth of information built over the last decade.**

***Index Terms*—Lexical resource, wordnet, multilingual interoperability, semantic web**

I. Introduction

ENACHING the development of multilingual lexicons is of foremost importance for intercultural collaboration to take place, as multilingual lexicons are the cornerstone of several multilingual applications (such as cross-language QA and IR, Machine Translation, terminology management, Multilingual computing, etc.). In addition, recently the availability of lexical resources and their multilanguage support has received growing attention by the Semantic Web community, as a rich and powerful mean that offers new possibilities to better handle and define the semantics of data. As a consequence, we have assisted to the first attempts of integration of lexical resource in the Semantic Web infrastructure and content organization model. Nevertheless, large-scale multilingual lexical resources are not as widely available and are very costly to construct.

The previous trend in lexical resource was oriented to maximization of effort by building large-scale, general-purpose lexicons. However, these lexical resources are not always satisfactory despite the tremendous amount of work needed to build them and the richness and degree of sophistication of the information contained therein; often lexical resources suffer an unbalanced coverage of their domain or are too much or too little detailed. Moreover, market calls for new types, rapidly built and easy tailored exploiting the richness of existing lexicons.

To meet these needs, lexical resources need to be made available, to be constantly accessed by different types of users, who may want to select different portions of the same resource, or may need to combine information coming from different resources.

This scenario no longer leaves space to static, closed, and locally managed repositories of lexical information; instead, it calls for an environment where lexical resources can be shared are reusable, and are openly customizable.

At the same time, as the history of the web teaches, it would be a mistake to create a central repository containing all the shared lexical resources because of the difficulties to manage it. Distribution of resources thus becomes a central concept: the solution proposed by the lexical resource community thus consists in moving towards distributed language *services*, based on open content interoperability standards, and made accessible to users via web-services technologies.

There is another, deeper argument in favor of distributed lexical resources: language resources, lexicons included, are inherently distributed because of the diversity of languages distributed over the world, that makes it impossible to have one single centralized repository of resources. In this way, each language resource is developed and maintained in its natural environment.

Having lexical resources available as web services would allow to create new resources on the basis of existing ones, to exchange and integrate information across repositories, and to compose new services on demand: an approach towards the development of an infrastructure built on top of the Internet in the form of distributed language services is presented in [1].

This new type of language resources can still be stored locally, but its maintenance and exploitation can be a matter of agents being choreographed to act over them.

Admittedly, this is a long-term scenario requiring the contribution of many different actors and initiatives (among which we only mention standardization, distribution and international cooperation). The first prerequisite for this scenario to take place is to ensure true interoperability among lexical resources, a goal that is long being addressed to by the standardization community and that is now mature.

Although the paradigm of distributed and interoperable lexical resources has largely been discussed and invoked, very little has been made in comparison for the development of new methods and techniques for its practical realization. Some initial steps are made to design frameworks enabling inter-lexica access, search, integration and operability. An example is the Lexus tool ([2]), based on the Lexical Markup Framework ([3]), that goes in the direction of managing the exchange of data among large-scale lexical resources. A similar tool, but more tailored to the collaborative creation of lexicons for endangered language, is SHAWEL ([4]). However, the general impression is that little has been made towards the development of new methods and techniques for attaining a concrete interoperability among lexical resources.

In this paper we present a tool, based on a web service architecture, fostering the integration and interoperability of computational lexicons, focusing on the particular case of mutual linking and cross-lingual enrichment of distributed monolingual lexical resources. As a case-study, we have chosen to work with two lexicons belonging to the WordNet family, the ItalWordNet [5] and Sinica BOW [6]. The development of this application is intended as a case-study and a test-bed for trying out needs and requirements posed by the challenge of semi-automatic integration and enrichment of practical, large-scale multilingual lexicons for use in computer applications.

The paper is organized as follows: section 2 describes the recent process of integration of wordnet in the Semantic Web, especially through the analysis of wordnet's World Wide Web Consortium (W3C) standard RDF/OWL representation; section 3 describes the general architectural design of our project; section 4 describes the tool taking care of cross-lingual integration of lexical resources, while a case-study involving an Italian and Chinese lexicons is presented in section 5. Section 6 briefly explains how this tool can be integrated in a more general framework for the semi-automatic management of lexical resources.

II. WordNet standard RDF/OWL representation: a data model for the Semantic Web Review Stage

During the last years, the lexical reference WordNet has received a growing attention by the Semantic Web research community. After the born of a 'WordNet Task Force' of the W3C's 'Semantic Web Best Practices Working Group' [SWBPWG] [7], WordNet has been translated in the widely adopted standard semantic languages RDF and OWL [8], and then has been published a Working Draft as a rielaboration and a synthesis of existing non-standard conversion.

RDF(S) and OWL, designed to describe collections of resources on the Web, are convenient data models to represent highly interconnected information and their semantic relations, and therefore useful to support WordNet graph data model. Moreover RDF/OWL representation of WordNet is easy extensible, allows for interoperability and makes no assumptions about a particular application domain.

The conversion is based on a hierarchy of classes and properties organized on the basis of the Princeton's WordNet Prolog distribution's conceptual structure. The reference's conceptual model has been changed only in the representation format, without affecting the original architecture.

WordNet model is composed by three main classes: Synset, WordSense and Word. The first two are divided into four

```
Synset
      AdjectiveSynset
              AdjectiveSatelliteSynset
      AdverbSynset
      NounSynset
      VerbSynset

WordSense
      AdjectiveWordSense
              AdjectiveSatelliteWordSense
      AdverbWordSense
      NounWordSense
      VerbWordSense
Word
              Collocation
```

Fig. 1. The class hierarchy of the WordNet RDF/OWL schema.

fundamental lexical types subsets: noun, verb, adjective and adverb. The only subset of Word is Collocation, used to represent words that have hyphens or underscores in them (Figure 1).

The properties:

1. represent lexical relations between the main classes, connecting couples of Synsets or WordSenses;
2. describe attributes of classes;
3. connect each Synset with WordSense/s (*wn:synsetContainWordSense*) and each

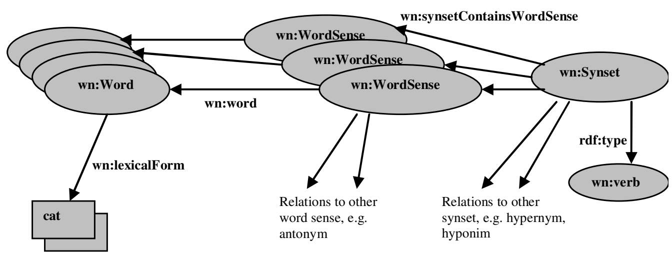

Fig. 2. Diagram of the schema of Wordnet RDF/OWL (prefixes 'wn' and 'rdf' stand for respectively the namespaces of WordNet and RDF(S)).

WordSense with the Word it represents (*wn:Word*).

Each Word is connected to its lexical form through the property *wn:lexicalFrom* and each Synset is characterized by a specific type (*rdf:type*) (Figure 2).

This representation of WordNet, composed of a single RDF/OWL schema, provides OWL semantics while still being interpretable by pure RDFS tools. Moreover, it defines a robust, human-readable URI assignment system, an on-line querying model based on the Common Bounded Description of resources and a reduced version of WordNet database (called WordNet Basic), so as to keep the footprint small when the complete set of relations is not needed.

The adaption of WordNet Web Services to support the RDF/OWL representation can represent another important step towards a stronger integration and an effective use of this important lexical resource into Semantic Web. This kind of resources could have a fundamental place in many Semantic Web base processes like ontology management, semantic interpretation of Web Services [9] and so on. Moreover, the future addition of interlingual information handling possibilities to RDF/OWL data model can support the achievement of real multiligual semantic interoperability in the Web. This topic has been explicitly left unsolved by W3C WordNet Task Force.

III. An Architecture for Integrating Lexical Resources

Designing a general architecture able to turn into reality the vision of shared and distributed lexical repositories is a very challenging task. We designed a distributed architecture to enable a rapid prototyping of cooperative applications for integrating lexical resources. This architecture is articulated in three layers (Figure 3):

- The lower layer consists of a sort of *meta-wordnet*, i.e. a grid of local wordnets realized as a virtual repository of generic XML or RDF/OWL databases residing at different locations and accessible through web services. Basic software services are also necessary, such as an UDDI server for the registration of the local wordnets and web services dedicated to the coherent management of the different versions of WordNet the databases refer to.
- The middle layer hosts several applications that exploit the wordnets grid. The so-called MultiWordNet Service (MWS, Section 3) was built as a proof of concept of the possibility to mutually enrich wordnets in a distributed environment; other, more advanced NLP applications (in particular multilingual) can be developed by exploiting the availability of the WordNet grid.
- A higher layer, called "cooperative layer" or LeXFlow is intended as an overall environment where all the modules realized in the lower layers are integrated in a comprehensive workflow of human and software agents.

In Section V we illustrate how the general LeXFlow environment could accommodate the tool described as a

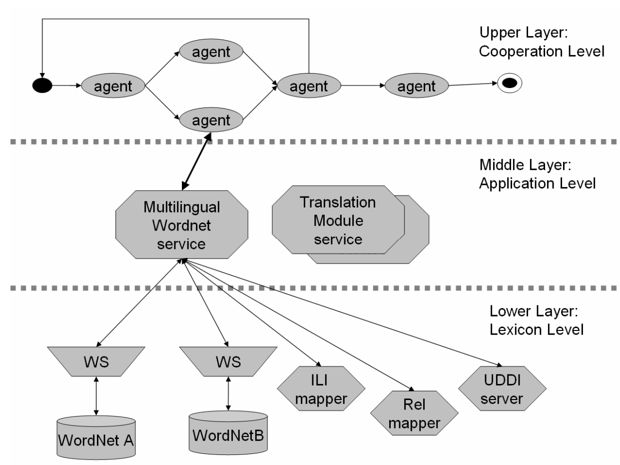

Fig. 3. A three-layered architecture for integrating lexical resources.

module of a general architecture geared towards lexicon management.

IV. Multilingual WordNet Service

In this section we present a tool that addresses the issue of lexicon augmentation or enrichment focusing on mutual enrichment of two wordnets.

This module, named "Multilingual WordNet Service" is responsible for the *automatic cross-lingual fertilization* of lexicons having a WordNet-like structure. Put it very simply, the idea behind this module is that a monolingual wordnet can be enriched by accessing the semantic information encoded in corresponding entries of other monolingual wordnets.

Since each entry in the monolingual lexicons is linked to the Interlingual Index (ILI, cf. Section 3.1), a synset of a WN(A) is indirectly linked to another synset in another WN(B). On the basis of this correspondence, a synset(A) can be enriched by importing the relations that the corresponding synset(B) holds with other synsets(B), and vice-versa. Moreover, the enrichment of WN(A) will not only import the relations found in WN(B), but it will also propose target synsets in the language(A) on the basis of those found in language(B).

The various WN lexicons reside over distributed servers and can be queried through web service interfaces. The overall architecture for multilingual wordnet service is depicted in Figure 4.

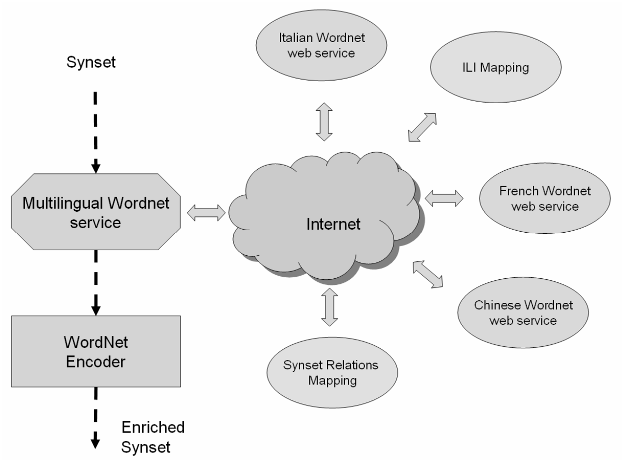

Fig. 4. Multilingual Wordnet Service Architecture.

Put in the framework of the general LeXFlow architecture, the Multilingual wordnet Service can be seen as an additional external software agent that can be added to the augmentation workflow or included in other types of lexical flows.

A. *Linking Lexicons through the ILI*

The entire mechanism of the Multilingual WN Service is based on the exploitation of Interlingual Index ([10]), an unstructured version of WordNet used in EuroWordNet ([11]) to link wordnets of different languages; each synset in the language-specific wordnet is linked to at least one record of the ILI by means of a set of equivalence relations (among which the most important is the EQ_SYNONYM, that expresses a total, perfect equivalence between two synsets).

In the schema of a WN lexical entry, under the root "synset" we find both internal relations ("synset relations") and ILI Relations, which link to ILI synsets.

Figure 5 shows the role played by the ILI as set of pivot nodes allowing the linkage between concepts belonging to different wordnets.

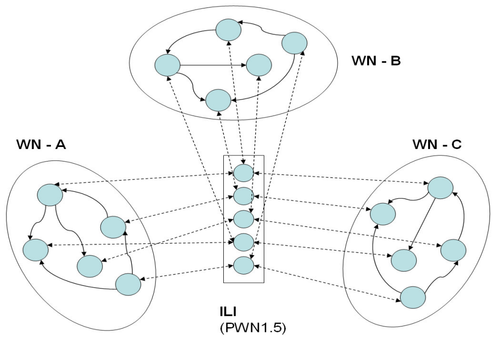

Fig. 5. Interlingual Linking of Language-specific Synsets.

In the Multilingual WN Service, only equivalence relations of type EQ_SYNONYM and EQ_NEAR_SYNONYM have been taken into account, being them the ones used to represent a translation of concepts and also because they are the most exploited (for example, in IWN, they cover about the 60% of the encoded equivalence relations). The EQ_SYNONYM relation is used to realize the one-to-one mapping between the language-specific synset and the ILI, while multiple EQ_NEAR_SYNONYM relations (because of their nature) might be encoded to link a single language-specific synset to more than one ILI record. In Figure 6 we represented the possible relevant combinations of equivalence relations that can realize the mapping between synsets belonging to two languages. In all the four cases, a synset "a" is linked via the ILI record to a synset "b" but a specific procedure has been foreseen in order to calculate different "plausibility scores" to each situation. The procedure relies on different rates assigned to the two equivalence relations (rate "1" to EQ_NEAR_SYNONYM relation and rate "0" to the EQ_SYNONYM). In this way we can distinguish the four cases by assigning respectively a weight of "0", "1", "1" and "2".

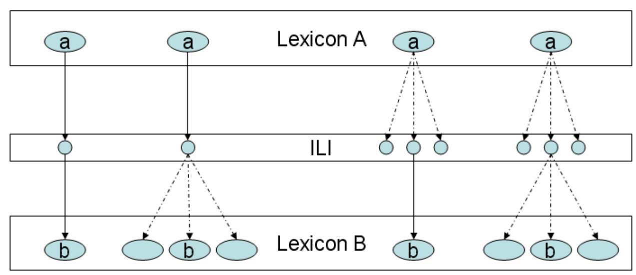

Fig. 6. Possible combinations of relations between Lexicons A, B, and the ILI.

The ILI is a quite powerful yet simple method to link concepts across the many lexicons belonging to the *WordNet-family*. Unfortunately, no version of the ILI can be considered a standard and often the various lexicons exploit different version of WordNet as ILI. This is a problem that is handled at web-service level, by incorporating the conversion tables provided by ([12]). In this way, the use of different versions of WN does not have to be taken into consideration by the user who ac-cesses the system but it is something that is resolved by the system itself .

B. *Description of the Procedure*

On the basis of ILI linking, a synset can be enriched by importing the relations contained in the corresponding synsets belonging to another wordnet.

In the procedure adopted, the enrichment is performed on a synset-by-synset basis. In other words, a certain synset is selected from a wordnet resource, say WN(A). The cross-lingual module identifies the corresponding ILI synset, on the basis of the information encoded in the synset. It then sends a query to the WN(B) web service providing the ID of ILI synset together with the ILI version of the starting WN. The WN(B) web service returns the synset(s) corresponding to the WN(A) synset, together with reliability scores. If WN(B) is based on a different ILI version, it can carry out the mapping between ILI versions (for instance by querying the ILI mapping web service). The cross-lingual module then analyzes the synset relations encoded in the WN(B) synset and for each of them creates a new synset relation for the WN(A) synset.

If the queried wordnets do not use the same set of synset relations, the module must take care of the mapping between different relation sets. In our case-study no mapping was needed, since the two sets were completely equivalent.

Each new relation is obtained by substituting the target WN(B) synset with the corresponding synset WN(A), which again is found by querying back the WN(A) web service (all these steps through the ILI). The procedure is formally defined by the following formula:

Let $a_j \in A$
Let $Ba_j=\{b_i \mid b_i \in B \text{ and } (b_i \text{ ILI } a_j)\}$
  $\forall\ b_i \in Ba_j$
  Let $R_i=\{b_i r_k b_p \mid b_i, b_p \in B \text{ and } (r_k \in R_A \cap R_B)\}$
  $\forall\ b_i r_k b_p \in R_i$
    Let $Ab_p=\{a_i \mid a_i \in A \text{ and } (a_i \text{ ILI } b_p)\}$
    $\forall\ a_t \in Ab_p$
      $a_j r_k a_t$ is a candidate relation

Legenda:
A,B lexicons
a_j, b_i synsets
$a_j r_p a_i$ synset relation r_p between a_j and a_i
$b_i \text{ILI} a_j$ b_i is connected by ILI with a_j
R_A, R_B relation space of lexicons B
$R_A \cap R_B$ the common relation space of B and A

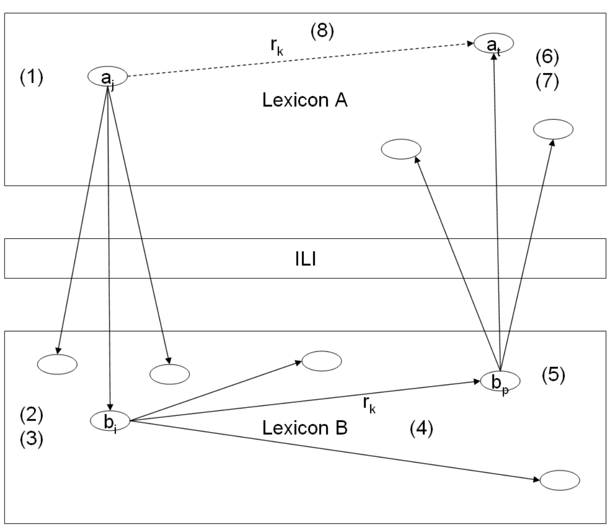

Fig. 7. A procedure for finding new relations.

V. A Case Study: Cross-fertilization between Italian and Chinese Wordnets.

We explore this idea with a case-study involving the ItalWordNet ([5]) and the Academia Sinica Bilingual Ontological Wordnet (Sinica BOW, [6]).

The BOW integrates three resources: WordNet, English-Chinese Translation Equivalents Database (ECTED), and SUMO (Suggested Upper Merged Ontology). With the integration of these three key resources, Sinica BOW functions both as an English-Chinese bi-lingual wordnet and a bilingual lexical access to SUMO. Sinica Bow currently has two bilingual versions, corresponding to WordNet 1.6. and 1.7. Based on these bootstrapped versions, a Chinese Wordnet (CWN, [13]) is under construction with handcrafted senses and lexical semantic relations. For the current experiment, we have used the version linking to WordNet 1.6.

ItalWordNet was realized as an extension of the Italian component of EuroWordNet. It comprises a general component consisting of about 50,000 synsets and terminological wordnets linked to the generic wordnet by means of a specific set of relations. Each synset of ItalWordNet is linked to the Interlingual-Index (ILI).

The two lexicons refer to different versions of the ILI (1.5 for IWN and 1.6 for BOW), thus making it necessary to provide a mapping between the two versions. On the other hand, no mapping is necessary for the set of synset relations used, since both of them adopt the same set.

For the purposes of evaluating the cross-lingual module, we have developed two web-services for managing a subset of the two resources.

The following Figure shows a very simple example where our procedure discovers and proposes a new meronymy relation for the Italian synset {passaggio,strada,via}. This synset is equivalent to the ILI "road,route" that is ILI-connected with BOW synset "道路,道 ,路" (dao_lu, dao, lu) (Figure 8, A) . The Chinese synset has a meronymy relation with the synset "十字路口" (wan) (B). This last synset is equivalent to the ILI "bend, crook, turn" that is ILI-connected with Italian WordNet synset "curvatura, svolta, curva" (C). Therefore the procedure will propose a new candidate meronymy relation between the two Italian WordNet synsets

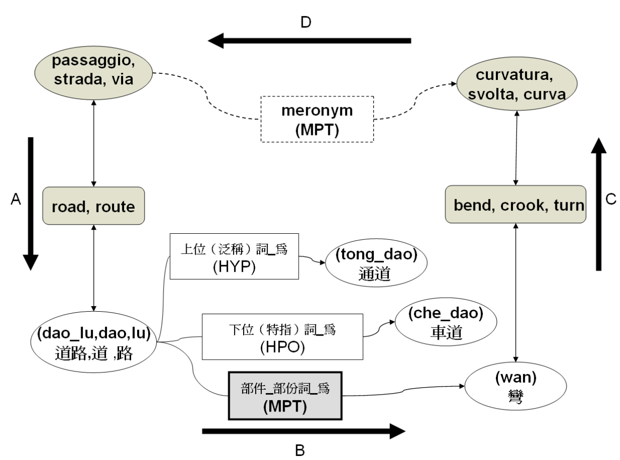

Fig. 8. Example of a new proposed meronymy relation for Italian.

(D).

Similarly, Figure 9 shows the flow of information between the two WordNets.

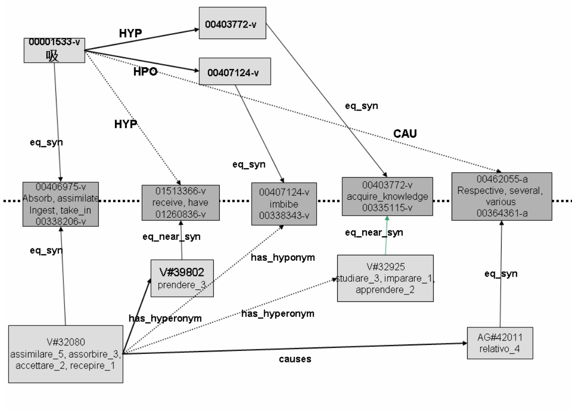

Fig. 9. Inferred relations for Italian and Chinese.

A. *Considerations and Lessons Learned*

Given the diversity of the languages for which wordnets exist, we note that it is difficult to implement an operational standard across all typologically different languages. Work on enriching and merging multilingual resources presupposes that the resources involved are all encoded with the same standard. However, even with the best efforts of the NLP community, there are only a small number of language resources encoded in any given standard. In the current work, we presuppose a de-facto standard, i.e. a shared and conventionalized architecture, the WordNet one. Since the WordNet framework is both conventionalized and widely followed, our system is able to rely on it without resorting to a more substantial and comprehensive standard. In the case, for instance, of integration of lexicons with different underlying linguistic models, the availability of the MILE ([14]) was an essential prerequisite of our work. Nevertheless, even from the perspective of the same model, a certain degree of standardization is required, at least at the format level.

From a more general point of view, and even from the perspective of a limited experiment such as the one described in this paper, we must note that the realization of the new vision of distributed and interoperable language resources is strictly intertwined with at least two prerequisites. On the one side, the language resources need to be available over the web; on the other, the language resource community will have to reconsider current distribution policies, and to investigate the possibility of developing an "Open Source" concept for LRs.

VI. LeXFlow

This MWNS can run as an individual system, but it has to be seen more as a software module to be integrated into the general LeXFlow architecture ([15]), developed with the aim to make the vision of an infrastructure for access and sharing of linguistic resources more tangible.

LeXFlow was born as an adaptation to computational lexicons of XFlow, a cooperative web application for the management of document workflows (DW, [16]) and can be considered as both an architecture for proving new cooperation methods among lexicon experts and a general, versatile framework enabling automatic lexical resource integration. The novelty of LeXFlow is that it enables the cooperation of agents, either human or software agents and allows different agents to interact, even residing over distributed places. Since it allows the independent and coordinated sharing of actions over portions of lexicons, LeXFlow naturally lends itself as a tool for the management of distributed lexical resources.

In the LeXFlow framework the workflow of lexical entries is described by a new XML application called XFlowML (XFlow Markup Language), largely based on XSLT Processing Model. XFlowML describes a workflow using an agent-based approach. Each human or software agent can participate to the workflow with one or more roles, defined as XPath expressions, based on a hierarchical role chart. An XFlowML document contains as many templates as are the agent roles participating in the workflow. The selection of the templates will establish the order with which the agents will receive the lexical entry. The document workflow engine constitutes the runtime execution support for the document processing by implementing the XFlowML constructs. To this end, at first we have defined the logical schema of a lexical entry and the contextual domain of the document workflow including all human and software agents cooperating, with different roles, to the compilation of lexical entries.

Finally we have formalized the procedural rules and the access control rules (XFlowML) of lexical entry compilation.

VII. Conclusion

Our work can be proposed as a prototype of a web application that would support the Global WordNet Grid initiative (www.globalwordnet.org/gwa/gwa_grid.htm).

Any multilingual process, such as cross-lingual information retrieval, must involve both resources and tools in a specific language and language pairs. For instance, a multilingual query given in Italian but intended for querying English, Chinese, French, German, and Russian texts, can be sent to five different nodes on the Grid for query expansion, as well as performing the query itself. In this way, language specific query techniques can be applied in parallel to achieve best results that can be integrated in the future. As multilingualism clearly becomes one of the major challenges of the future of web-based knowledge engineering, WordNet emerges as one leading candidate for a shared platform for representing a lexical knowledge model for different languages of the world. This is true even if it has to be recognized that the wordnet model is lacking in some important semantic information (like, for instance, a way to represent the semantic predicate). However, such knowledge and resources are distributed. In order to create a shared multi-lingual knowledge base for cross-lingual processing based on these distributed resources, an initiative to create a grid-like structure has been recently proposed and promoted by the Global WordNet Association, but until now has remained a wishful thinking. The success of this initiative will depend on whether there will be tools to access and manipulate the rich internal semantic structure of distributed multi-lingual WordNets. We believe that our work on LeXFlow offers such a tool to provide interoperable web-services to access distributed multilingual WordNets on the grid. This allows us to exploit in a cross-lingual framework the wealth of monolingual lexical information built in the last decade.

In conclusion, in this effort to reach multilingual interoperability using wordnet, we must also consider how the growing adoption of wordnet in the Semantic Web community and its integration with Semantic Web technologies and data models, previously described (see section 2), could attract more and more interest on the process of standardization and wide multilingual availability and interoperability of this high-quality lexical resource.

REFERENCES

[1] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), keynote address (2006) 96-100.

[2] Kemps-Snijders, M.. Nederhof, M., Wittenburg, P.: LEXUS, a web-based tool for manipulating lexical resources. In Proceedings of LREC2006. Genoa, Italy (2006) 1862-1865

[3] Francopoulo, G., George, M., Calzolari, N., Monachini, M., Bel, N., Pet, M., Soria, C.: Lexical Markup Framework (LMF). In Proceedings of LREC2006. Genoa, Italy (2006) 233-236

[4] Gulrajani, G., Harrison, D.: SHAWEL: Sharable and Interactive Web-Lexicons. In Proceedings of the LREC2002 Workshop on Tools and Resources in Field Linguistics. Las Palmas, Canary Islands, Spain (2002) 1-4

[5] Roventini, A., Alonge, A., Bertagna, F., Calzolari, N., Girardi, C., Magnini, B., Marinelli, R., Zampolli, A.:. ItalWordNet: Building a Large Semantic Database for the Automatic Treatment of Italian. In Antonio Zampolli, Nicoletta Calzolari, and Laura Cignoni, editors, Computational Linguistics in Pisa, IEPI, Pisa-Roma (2003) 745-791

[6] Huang, C., Chang, R., Lee, S.: Sinica BOW (Bilingual Ontological Wordnet): Integration of Bilingual WordNet and SUMO. In Proceedings of LREC2004. Lisbon, Portugal (2004) 1553-1556

[7] World Wide Web Consortium 'Semantic Web Best Practices Working Group' – Web Site http://www.w3.org/2001/sw/BestPractices/

[8] World Wide Web Consortium – RDF/OWL Representation of Wordnet - W3C Working Draft 19 June 2006 - http://www.w3.org/TR/wordnet-rdf

[9] Luke Simon, Ajay Bansal, Ajay Mallya, and Gopal, Thomas D. Hite - A Universal Service-Semantics Description Language – Gupta, Department of Computer Science, University of Texas at Dallas (2005).

[10] Peters, W., Vossen, P., Diez-Orzas, P., Adriaens, G.: Cross-linguistic Alignment of Wordnets with an Inter-Lingual-Index. In Ide, N., Greenstein, D., Vossen, P. (eds.): Special Issue on EuroWordNet. Computers and the Humanities 32(2-3) (1998) : 221-251

[11] Vossen, P.: Introduction to EuroWordNet. In Ide, N., Greenstein, D., Vossen, P. (eds.): Special Issue on EuroWordNet. Computers and the Humanities 32(2-3) (1998) 73-89

[12] Daudé, J., Padró, L., Rigau, G.: A Complete WN1.5 to WN1.6 Mapping. In Proceedings of NAACL Workshop "WordNet and Other Lexical Resources: Applications, Extensions and Customizations". Association for Computational Linguistics, Pittsburg, PA, USA (2001) 83-88

[13] Huang, C., Chen, C., Weng, C., Lee, H., Chen, Y., Chen, K.: The Sinica Sense Management System: Design and Implementation. Computational Linguistics and Chinese Language Processing 10(4) (2005) 417-430

[14] Calzolari, N., Bertagna F., Lenci, A., Monachini, M. (eds.): Standards and Best Practice for Multilingual Computational Lexicons. MILE (the Multilingual ISLE Lexical Entry). ISLE CLWG Deliverable D2.2 & 3.2. Pisa. (2003)

[15] Ruimy, N., Monachini, M., Gola, E., Calzolari, N., Del Fiorentino, C., Ulivieri, M., Rossi, S.: A Computational Semantic Lexicon of Italian: SIMPLE. In Zampolli, A., Calzolari, N., Cignoni, L. (eds.): Computational Linguistics in Pisa. IEPI, Pisa-Roma (2003) 821-864

[16] Marchetti, A., Tesconi, M., Minutoli, S.: XFlow: An XML-Based Document-Centric Workflow. In Proceedings of WISE'05. New York, NY, USA (2005) 290-303